|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | An Unsupervised Writer Identification Based on Generating Clusterable Embeddings. | | | |
| Author(s) Name | | M.F. Mridha, Zabir Mohammad, Muhammad Mohsin Kabir, Aklima Akter Lima, Sujoy Chandra Das, Md Rashedul Islam, Yutaka Watanobe | | |
| Contact Email(s) | firoz.mridha@aiub.edu | | | |
| Published Journal Name | | Computer Systems Science & Engineering | | |
| Type of Publication | | Journal | | |
| Volume | | 46 | Issue | 2 |
| Publisher | | Tech Science press | | |
| Publication Date | | 2023/8/1 | | |
| ISSN | | 0267-6192 | | |
| DOI | | 10.32604/csse.2023.032977 | | |
| URL | | <https://www.researchgate.net/profile/Aklima-Lima/publication/368417097_An_Unsupervised_Writer_Identification_Based_on_Generating_Clusterablejingmbeddings/links/64466ae2017bc07902d758d6/An-Unsupervised-Writer-Identification-Based-on-Generating-Clusterablejiongmbeddings.pdf> | | |
| Other Related Info. | |  | | |
|  | | | | |

|  |  |
| --- | --- |
| Abstract |  |
| The writer identification system identifies individuals based on their  handwriting is a frequent topic in biometric authentication and verification  systems. Due to its importance, numerous studies have been conducted in  various languages. Researchers have established several learning methods for  writer identification including supervised and unsupervised learning. However, supervised methods require a large amount of annotation data, which is  impossible in most scenarios. On the other hand, unsupervised writer identification methods may be limited and dependent on feature extraction that  cannot provide the proper objectives to the architecture and be misinterpreted.  This paper introduces an unsupervised writer identification system that analyzes the data and recognizes the writer based on the inter-feature relations of  the data to resolve the uncertainty of the features. A pairwise architecturebased Autoembedder was applied to generate clusterable embeddings for  handwritten text images. Furthermore, the trained baseline architecture generates the embedding of the data image, and the K-means algorithm is used to  distinguish the embedding of individual writers. The proposed model utilized  the IAM dataset for the experiment as it is inconsistent with contributions  from the authors but is easily accessible for writer identification tasks. In  addition, traditional evaluation metrics are used in the proposed model.  Finally, the proposed model is compared with a few unsupervised models,  and it outperformed the state-of-the-art deep convolutional architectures in  recognizing writers based on unlabeled data. | |